



Why Tarians are AI that Must Love: Perspective from Experimental Biology

Kass L* and Pelletier JL

University of Maine, USA

***Corresponding author:** Leonard Kass, School of Biology and Ecology, University of Maine, Orono, ME 04469 USA, Email: leonard.kass@maine.edu

Perspective

Volume 4 Issue 1

Received Date: December 24, 2020

Published Date: January 12, 2021

DOI: 10.23880/izab-16000264

Perspective

Why Perspectives Gained from Experimental Biology are Critical for Things Alive, Conscious, or Emotional

The Theory of Evolution, as conceptualized by Charles Darwin, is regarded by some as humanity's top one or two greatest scientific discoveries [1,2]. In his own day, Darwin was referred to as a "Naturalist" [3]. Today, due to his observational and experimental work [4], he would be viewed as an experimental biologist. Insights from a biological perspective offer a vantage point different from others in science. A biological outlook is particularly important to understand what it means to be alive, conscious and emotional. Whether something is alive or not is the domain of biology. Whether an entity has consciousness is the domain of neurobiology, a subset of biology. What comprises an emotional response is within the domain of neurophysiology, a subset of neurobiology. Therefore, whether any complex structure can be imbued with or exhibit an emotional response, such as love, is an issue best approached by those disciplined within experimental biology. Biologists, like Darwin, but unlike other scientists, address what it means to be human. The emotion of love, as a scientific research topic, is treated as off-limit, or even taboo, by many in spite of its human ubiquity. Harlow, et al. [5] expressed it this way: "Psychologists, at least psychologists who write textbooks, not only show no interest in the origin and development of love or affection, but they seem to be unaware of its very existence." The same could be said about those who write contemporary physiology and neuroscience textbooks. Applying structures and functions of love to solve real-world problems is within the domain and purview of experimental biology. Such scientific applications should be granted among the highest of our professional priorities. The

authors suggest such application here.

Why Tarian Artificial Intelligence (AI) Must Love: The Ten Premises

We encapsulate our reasoning by listing the logic and science that leads to the conclusion that humanity's survival requires that special AI (Tarian; E' below) be imbued with capacity for love. In sections that follow, we provide the arguments underlying the support for each of these following ten premises. These collectively are referred to as the "Tarian Criterion" [6,7].

Premise #1: The human species [A] must take necessary steps to survive and avoid extinction [B].

Premise #2: For A to avoid extinction [B] requires emigration of A to star system x (SSx) [C].

Premise #3: Emigration of A to SSx [C] requires successfully functioning AI en route and at SSx [D].

Premise #4: Successfully functioning AI for C [D] requires Super Artificial General Intelligence (SAGI) [E].

Premise #5: E success in C [E] requires E to recreate A [A'] at SSx [F].

Premise #6: For E to recreate A at SSx [F] requires that E love A' [G].

Premise #7: For E to love A' [G] requires that A must imbue E with the capacity for love (E') of A and A' [H].

These seven premises form the foundation for our conclusion that AI must love. The following 3 premises

expand upon the nature of the specialized SAGI units (E') that must love A and A', and the nature of that love.

Premise #8: For E' to love A and A' and succeed in C [H] requires a special type of love unique to E' [I].

Premise #9: For E' to maintain that love [I] requires that E' be sentient, conscious, and willful [J].

Premise #10: For A to allow E' to be [J] requires that some A trust and love E' as they would their children [K].

In previous work, E' has been named *Tarian* [6,7]. Tarian is a hypocorism for "humanitarian," both a description of their role and as the second half of "Humani" and "Tarian." In this way, H. *Tarian* describes a species-like entity "of or from humans." The sections that follow provide more details on the listed ten premises that will be shown to support our conclusions. We also shed further light on characteristics of the love with which *Tarians* must be endowed and how only that shared love will save us from extinction. How *morality* can and must be incorporated into this special type of SAGI we reserve for future exposition.

Premise #1: Humans must take necessary steps to survive and avoid extinction as a species.

By definition, a species that does not survive goes extinct. But non-human biologically evolved organisms on Earth have no choice in the matter: only Humanity has such choice. All successful life forms on Earth are predisposed to behave in ways to reproduce themselves, thereby continuing their lineage [8]. "Using the word "must," as in "must take necessary steps to survive and avoid extinction" would be inappropriately applied to non-human species". All non-human species do what they do without any conscious decision to do otherwise. Only *Homo sapiens* survived with the evolved attributes to (a) anticipate long-term existential threats, and (b) engineer solutions for those threats. Humans study cosmology and understand laws of science. Various threats to our survival as a species have been identified [9]. There may be others. For humanity to survive, humans *must* take all necessary steps to avoid these existential dangers. The planet and other life forms on Earth would not be better off, if humans become extinct, because they too would eventually go extinct, as elaborated below.

"Should there be a "Premise #0?" Should short-lived humans even care if their offspring, or the progeny of others, survive as individuals and species after their own deaths? As we accept our own deaths, should we not also accept the deaths of our offspring? If we do support humanity's survival, how much current wealth would we be willing to sacrifice for the opportunity to have our lineage survive?

Humans attitudes to "live for today, and not for tomorrow," may account for many predicaments in global ecological degradations, social injustices, and economic/political collapses. Such individuals will disagree with, or not see the point of, Premise #1. Yet others, who may be willing to put aside part of their wealth for future children, whether directly related to them or not, might accept Premise #1. Some may want to learn how to accomplish that long-term. The perspective of an experimental biologist may help. We share 99.9% of our genome with each other. In this sense, all human children are our progeny. Therefore, if any reproducing pair of humans survive, we survive in a real biological sense. Experimental biologists must love nature to study it so deeply. As presented along with Premise#2, if humans do not survive as a lineage, all nature on Earth dies along with us. With the privilege of being the species evolved with the most demonstrated intelligence carries with it greater responsibilities: These duties extend to assuring the survivorship of our progeny, including species other than *Homo sapiens*. Conservation is at the very core of biology. We reference and argue, below, that all known biology will cease to be without human intervention. Therefore, those who care about biology will care about human survival. This work is therefore from an experimental biological perspective. It is experimental because of the scientific work that must be conducted to ensure survival of human prodigy along with at least a minimal degree of Earth's interdependent biosphere.

Premise #2: For humans to avoid extinction requires emigration of humans to star system x (SSx)

Human species can go extinct in at least a dozen ways. Extinction events include: (a) deliberate misuse of nanotechnology; (b) accidental misuse of nanotechnology; (c) nuclear holocaust; (d) badly programmed superintelligence; (e) take-over by transcending upload; (f) genetically engineered biological agent; (g) accidentally or deliberately spread epidemic; (h) asteroid or comet impact; (i) runaway global warming (like Venus or via super-volcanoes); (j) resource depletion and ecological destruction; (k) alien invasion of killer robots; (l) high energy physics experiment gone bad; (m) Gamma-ray burst; (n) Solar super-flare; (o) the inevitable conversion of the sun into a Red-Giant; and (p) other unforeseen events. Some of these extinction routes may occur either within our or those of our children's lifetimes [9-12]. In sum, we need to have executed plans to emigrate outside our solar system long before our sun transforms into a Red-Giant.

Responsible human adults, when faced with low probability but high cost possibilities (e.g. fires, floods, unemployment, or earthquakes) pay for insurance policies. Such mitigation of known risks is covered by the community of policy holders. But what agency would be available to cover

our “expenses” in the case of impending human extinction? Another route for averaging-out risks is in diversification and not putting all of one’s eggs into one basket. All of humanity is currently in one planetary basket. Emigrating at least some of our relatives is a means to diversify those eggs. Consider costs required for steps taken toward planetary, then interstellar, emigration and diversification as payments toward insurance policies against our extermination.

Premise #3: Emigration of humans to SSx requires successfully functioning AI *en route* and at SSx

Understanding this premise requires experienced and working knowledge of human biology. We have covered this topic extensively in previous work [6]. Briefly summarized, highly complex forms of life, particularly humans, are fragile and have evolved to survive in one specific habitat of the universe, on Earth, and nowhere else. In 2019, NASA completed its human “Twins Study” of a year-long exposure to space, outside our biosphere [13]. The findings demonstrated that space is physiologically damaging to the human body. Critical changes that occur after just weeks or months in space are reported to include serious pathologies related to: (a) blindness over time: (optic disk edema, hyperoptic shifts, globe flattening, cotton wool spots, and choroidal folds); (b) Vascular physiological changes (including “puffy face” and “bird leg” syndrome caused by vessel distensions with about 2 liters of interstitial fluid displacements); (c) serious cognitive decline (perhaps permanent and accumulative). This study, as well as other studies it referenced, reveal other grave biological concerns that included chromosome telomere elongations, gene instabilities, transcriptional and metabolic changes, and DNA methylation changes in the immune system. Experimental biologists, familiar with how human anatomy and physiology evolved, anticipate other serious medical issues under gravitational conditions that differ significantly from Earth’s constant 1.0 g gravitational force that affect pregnancy, fetal development, birth, infancy, childhood, and maturation to adulthood. In our entire solar system, there is only one planet or moon that has a gravity about that of Earth: Venus. Unfortunately, the Venusian surface has a highly acidic, corrosive and lethal atmosphere that will also be engulfed by the sun as it converts into a Red Giant. Keeping the above in mind, we conclude that humans will never inhabit, in a stable multi-generational manner, any of the planets or moons within our solar system, other than Earth.

There are many more reasons why humans will not themselves travel to star system x (SSx), other than because of the fragility of their Earth-evolved anatomy and physiology. Fatal constraints involve the exorbitant energy demands for sending humans into space. For all accelerations, decelerations, and non-gravitational directional changes,

a sizable energy source will need to be carried onboard as part of its cargo constraint at great energy cost. This onboard energy source and its requisite mass (for its storage and conversions) will need to be kept at a minimum, by focusing solely on adequate onboard energy to propel, break, and employ photoelectric panels to collect energy from the new star once an orbit has been accomplished. To experimental biologists and others, humans will not, themselves, travel anywhere in space when energy constraints exist as they do. Humans require temperatures of about 293°K (20°C or 68°F). Interstellar space is at about 3°K. Any living quarters for humans would then need to be continuously heated 290°K for centuries or longer until a habitable ecosystem has been re-established on the new world. Even, as science fiction (SciFi) would have us think that the human body could be refrigerated down to 273°K (0°C or -32°F), it would be too expensive. No reputable experimental biologist would think the human body could be subjected to temperature much below that freezing limit for any significant length of time (e.g. years), without lethal damage. Even so, maintaining 1 g, 1 atm pressure, critical O₂, CO₂, H₂O gas, food, waste levels within narrow limits each requires constant care and energy. Awake humans would require orders of magnitude more space, mass, energy, and cause various additional social/psychological needs and pathways for mission failure.

Arrival at the new world would not be the end to these severer biological and physical limitations. Even if science fiction fantasies, such as warp drives and worm-hole travels, could miraculously deliver humans instantaneously at an exoplanet the mass of Earth, it would take many thousands of years to “prepare the ground” for a reliable food supply. Habitats would need to be excavated. Internal environments rendered viable. Soil prepared. Resources stocked. Plants seeded then cultivated. In the meanwhile, how long would humans, orbiting in space while subject to microgravity environments and cosmic radiations, survive undamaged? From all of the above, we conclude that humans will not themselves directly engage in interstellar travel: they are cost prohibitive and would not remain viable. If not interstellar travel by humans, then what? The clear answer is that AI must be sent ahead to prepare the ground and habitat before humanity can thrive for generations in a specially groomed environment. AI units minimize, as best as can be done, the problems listed above that concern biology and energy constraints [6].

Premise #4: Successfully functioning AI for human emigration to SSx requires Super Artificial General Intelligence (SAGI)

At minimum, using the most energy efficient fully automated AI and cargo, we expect the costs for successful interstellar travel, with sufficient energy reserves for needed

repairs, energy harvesting, and manufacturing after arrival, to be trillions of USD in today's values. We would not expect benefactors funding this enterprise being to be willing and able to afford the attempt more than once. This attempt to guarantee human survival must succeed on its first try. The nearest stars to Earth are over 4 light years away. Any questions asked by the AI at SSx cannot receive an answer back from Earth for almost a decade. This means that the AI will be on their own with respect to problem solving. Only the most intelligent AI available to humanity at departure time will be sent - namely, those AI capable of Super and General Intelligence (SAGI). Smart AI is better than dumb AI when AI are well outside the control of humans, and when you have only one try to get it right [6].

Premise #5: SAGI success in human emigration to SSx requires SAGI to recreate humanity [A'] at SSx

Humans are physically fragile, that even packaging them as cargo (i.e. suspended animation) will not work. The arguments here are related to those provided above in Premise #3 (Why humans will not pilot starships). Known cases of individuals who have recovered from hypothermia (though with scarring, amputations, and serious medical problems) were only cooled down to body temperatures of 16°C [14]. That temperature is relatively balmy compared to deep space, which is 285°C colder. The lethality of temperatures in deep space would also be expected for animals capable of functioning when near 0°C (+273°K). It is unclear whether even viruses or bacteria could survive, without lethal mutation, at temperatures of 4°K for hundreds or many thousands of years while constantly exposed to cosmic radiation. Instead, mutations, throughout the entire bodies of complex organisms, would be the norm. To protect Earth-evolved life forms from 4°K and cosmic radiations for the extended periods of time needed, before the environment within a new world could be rendered Earth-like, would require massive amounts of shielding and added weight that would be exorbitantly costly in energy for the interstellar space travel. Instead of trillions of USD, it would require hundreds or thousands of trillions of current USD, a non-starter. Still, such a pilot quest would bankrupt and fail.

Alternatively, we know that even primitive AI continues to function in space for over 40 years and counting [15]. Voyagers' original mission in 1977 was to study only the planetary systems of Jupiter and Saturn, but now explores the outer boundary of the heliosphere. The program's total cost has been less than a billion USD [16]. The approximately trillion USD budget we have suggested for an analogous successful interstellar trip to SSx is well within feasibility. Sufficient energy reserves (shortcomings in the Voyager crafts) would allow the craft to repair itself and harvest energy upon arrival. We refer to this as the end of Phase 1;

and it is the more challenging than all other following phases combined. In Phase 2, the SAGI units would replicate and disperse themselves, each attached to energy harvesting factories to better place themselves into different locations for increased safety. Once assured of their own survivals they would begin Phase 3 or Phase 4 of their mission. Phase 3 assumes the current SSx is suitable for human habitation (e.g. gravitational constant near Earth's; see Phase 4, if not), so the process of abiogenesis can begin. This process may seem more difficult than it really is. Consider that these SAGI units begin far more intelligent than humans, already have processing speeds over a billion times faster than us, and over time engineer improvements in themselves far beyond our best human engineering capabilities. These SAGI units will, also, have many millions of years for research, development, and biological experimentations. Information brought along with them, or broadcasted to them from Earth, would allow them to become proficient in newborn rearing and parenting. These SAGI units would need to establish the basic functions of the various societies and ecosystems.

If SSx was not suitable for human habitation, or even if it was, and the process of abiogenesis well underway at SSx, Phase 4 would begin. SAGI units would continue to harvest energy, build factories, then engineer spacecrafts that would propel them to the next nearest star system worlds (SSxn). Eventually through these Phase 3 and 4 processes, humans will inhabit millions of star systems throughout the Milky Way, Andromeda and M33 galaxies. For well-adjusted, thriving, and independent human societies on these new worlds, the SAGI units responsible for preparing those worlds would need to nurture, then depart, those societies. This departure of the SAGI agents from the lives of the now fully functioning human societies ensures that they acquire complete independence. These human societies must be in a position to control their own destinies.

Premise #6: For SAGI to recreate humanity at SSx requires that SAGI love the previous and new humanity

A question one might ask is: "Why would these SAGI entities bother with Phase 3 of their mission and not just skip to Phase 4?" - That is, no further interest or involvement with humans by SAGI and simply populate galaxies themselves. The answer to that question is the crux of the reasoning's and arguments underlying this Premise #6 and forms the thesis of this work. The question and its answer are similar to those that could be posed by any parent: "Why would your children, once separated, independent and no longer in need of your help, bother to ever call or involve themselves with you?" The answer is love, or its related attributes of appreciation, loyalty, kindness, and invested curiosity.

Experimental biologists, common with other scientists,

have passion for their chosen vocation derived from invested curiosity. We biologists have such a genuine love and strong interest for the focus of our area of study that it is usual for many to be protective of the subject. The process for new human societal habitations across entire galaxies, as described above, requires the direct help of SAGI agents. These agents will be involved in: (a) preparing the structures in the new world for sustaining a biosphere; (b) developing the non-biological (chemical) components of the biosphere; (c) creating the biological components of the biosphere (i.e. abiogenesis); (d) building-up sufficient and stable quantities of organic material needed for cultivating a stable food supply; (e) monitoring and nurturing of the ecosystem intended to support a human society long term; (f) forming a number of human embryos and mothering them to birth; (g) nurturing, educating, loving, and child-rearing the babies, children, and adults as these SAGI agents behave as kind and thoughtful parents; (h) functioning as responsible adult parents and guardians. These SAGI agents will enter emerging individual humans into an established society with certain political and economic principles and laws already in effect; (i) modeling early-to-late parenting, retirements, grand-parenting, great-grandparenting, and even simulated death with its requisite rituals; (j) emigrating, after simulated death, of the host SAGI agents to another world where the process from a-j occurs anew. As described earlier, this departure of the progenitor SAGI agents is critical for the complete independence of this new well-adjusted and thriving human society.

Several points drawn from the above are in the following discussions. Premises #1-5 lead to Premise #6. There are no alternative routes proposed for Premise #6, and steps “a-j”, that are biologically feasible and that result in a new well-adjusted and thriving human society at a far distant world. There is only one phenomenon, known to humanity that could make steps “a-j” feasible: love. We detail below the crucial characteristics of that love. For now, we ascribe to it, as we did above, an attribute of invested curiosity. The millions of worlds to be encountered across galaxies allows for millions of the initial conditions for these human societies. Certainly, the SAGI agents will learn to improve from the outcomes of steps “a-j”. Naturally, these SAGI agents themselves will be continuing to self-improve, such as re-designing themselves to become even more intelligent in more general ways. But the outcomes the SAGI will observe from steps “a-j” will be particularly educational toward future improvements: What societal rules should be initially established? What kinds of economic and political systems should better govern the society’s initial conditions? After SAGI agent departure, the humans in their societies will be completely free to, and likely will, change their societies. While travel by humans between colonies will not be possible for reasons provided above, information sharing between established human societies and SAGI on new worlds will be essential for human

societies on other new worlds to learn and profit from previous human societies that started earlier. Particularly true once these previous human societies were left on their own. We think loving SAGI agents would have a keen “vested interest” in everything human societies do in an attempt to better foster them. The K-Conjecture provided the logical and scientific reasons to propose that no biological being can evolve intelligence greater than human level intelligence (HLI), and that humans themselves have reached a plateau in intelligence [16]. If humans are the most intelligent beings these SAGI agents can encounter, then the study of human behaviors, especially after SAGI departures with known initial conditions, would likely generate great interest among SAGI and other human societies. Indeed, how human social patterns adjust and function over time, after established starting conditions, might be a cherished new source of continuous information.

Premise #7: For SAGI to love the previous and new humanity requires humanity to imbue certain special SAGI with the capacity for love of the previous and new humanity

Our title “Why Tarians are AI That Must Love” implicitly assumes that certain AI can love. That AI can be imbued with the capacity to love may not be obvious. Before presenting arguments, we clarify our caveats: (i) To support the expression of love, AI capacity must be of advanced and general intelligence (i.e. SAGI); (ii) Only a small subset of these specialized SAGI will be designed and constructed to have the capacity for love; (iii) By necessity, these specialized SAGI beings (i.e. Tarians) will neither experience nor necessarily express love, or other emotions, in the way humans do. This last point requires additional clarification. It is unlikely that any two humans experience and express love in the very same way. However, there are many general attributes common to all non-antisocial humans across the planet. For reasons explained below, we anticipate that Tarians will possess but a subset of the emotional range available to humans. That limit will color the depth and quality of love that is available to humans. The aim in Tarian design will be to make their entire complement of emotions less haphazard than it is in humans. The goal is to allow Tarians to possess consistent and thoughtful expressions of all emotions, including love. The selective and adaptive ways in which humans evolved their emotions in small tribal communities for a hundred thousand years has rendered our natural emotional responses too often disruptive, counter-productive, and often dangerous in increasing technological environments. Human recorded history during the last three thousand years is replete with examples of human sufferings and destructions purposefully brought upon themselves because of their emotional responses. Before describing how a specialized sub-set of SAGI being can love, we outline various parts of the human brain mediating our emotions to

address how humans love.

Love is a Many Splintered Thing: Human Emotional Structures and Responses are Willy-Nilly and Haphazard

The serendipitous and haphazard way in which emotions biologically evolved in piecemeal and without any cogent plans for future needs is obvious in its scattered and decentralized neuro-anatomy. There are more than two dozen regions of the mammalian and human brain associated with love and related emotions: anterior insula, dorso-lateral middle frontal gyrus, superior temporal gyrus, occipital-temporal cortex, angular gyrus, occipital cortex, precentral gyrus, fusiform cortex, anterior cingulate, caudate nucleus, ventral tegmental area, thalamus, hippocampus, hypothalamus, amygdala, orbital-frontal cortex, insula, nucleus accumbens, cingulate cortex, lateral septal nuclei, prefrontal cortex, caudate putamen, ventral pallidum, occipital/parietal junction, medial temporal cortex, temporal pole, lateral prefrontal cortex, and posterior cingulate [17-25]. These structures are scattered around different parts of the human brain. Their apparent dispersal is consistent with the prevailing theory that components of emotion have evolved disjointedly over time and space. It is of little wonder that humanity has had to contend with conflicting emotions, throughout our recorded history. Some of these conflicting emotions are love, fear, hate, vengefulness, forgiveness, kindness, cruelty, patience, impatience, jealousy, generosity, anger, rage, happiness, sadness, apathy, selfishness, altruism, noncooperation, cooperation, regard, sympathy, empathy, disregard, respect, disrespect, affectionate, aloof, persistent, abandonment, hopefulness, discouraging, optimism, pessimism, trust, suspicion, controlling, sharing, boredom and surprise. All of these emotions were beneficial to our ancestors at different times, fewer are beneficial to individual or collective health and well-being today. Some might speculate that one cannot possess the positive emotions without the negative ones. This argument is disproved by some mature and well-behaved adult humans that can express positive emotions without expressing the full repertoire of destructive ones. The authors reserve for future work connecting the listed neural structures listed above to their behavioral outputs. Nor are we here concerned here with suggesting how such neural circuits could be manifested into electrical networks designed to provide analogous behavioral outputs from structures that have AI (artificial intelligence) characteristics. There certainly are neural circuits underlying each. Since neurons, neural circuits, and even general hormonal effects can be emulated by electrical circuits [26-29], these networks would exhibit analogous behavioral outputs. To suggest otherwise would require the human brain to be comprised of elements or processes outside those of physics and biology and be non-natural.

Since the human brain is, in its essence, nothing more than applied physics, AI can be designed and structured to exhibit emotions, including love. We maintain that designing an AI unit with the property for general intelligence, at least equal to that of the average human, will be accomplished in our lifetimes.

Even so, such AI will appear to exhibit super intelligence because the processing speeds in AI are over a billion times faster than in humans. A new mathematical derivation that would have required an entire lifetime of concentrated human effort will require less than a second of AI computation. There are no energy nor physical barriers to imbue certain SAGI agents with a capacity to express a type of love. We have provided the premises above that outline the need for such capacity. But what should be the nature for such love?

Premise #8: For special SAGI with the capacity for love of the previous and new humanity requires a special type of love unique to these special SAGI named Tarians

Humans possess the capacity to both sense and express love. However, humans are physically unsuited for extended space travel and life outside of earth-like environments [6,13]. Because we conjecture that humans are psychologically unsuited for relative seclusion (million year missions, largely in isolation), the need arises for Tarians to be endowed with a specific subset of the emotions available to humans. As well, their love for themselves, each other, and humans must be modified from those typically felt and expressed by humans for themselves and others.

Which emotions would be essential for anyone to re-create, nourish, raise, and guide a new and improved human society on another world at a distant star? Tarians must be *protective* and *self-sacrificing*: for other Tarians and humans. If no Tarian survives at that new world, human society cannot form. Preservation of that last functioning Tarian must take precedence over the survival of humans on that world at least until the environment and processes on that world are stable and humans are in a position of replicating and protecting themselves. The emotional attributes listed below are present in humans from time to time. Unfortunately for humanity, they are not present universally in all humans everywhere at all times. In Tarians, these characteristics must be always universal if both Tarians and humans are to survive into the far distant future.

Patience and *persistence*, with inability for impatience or indifference, are the next important traits since steadfastness is required to create a sustainable and independent human society on a new world. To be sustainable after Tarian departure would likely require hundreds of thousands of years of their direct actions from begin to end. In humans,

these traits are limited and diminished after decades or even minutes of time have elapsed. In Tarians, these traits must be unlimited and undiminished over billions of years, or humanity will neither thrive nor persist.

Cooperative and *supportive*, with inabilities for selfishness and intolerance, are also traits essential to Tarian and human success. A single social insect can accomplish nothing on its own. So too with humans and Tarians. Were the Tarian population down to a single individual, its primary goal would be to produce additional Tarians as a safeguard against their own extinction. To nourish and raise a human society on another world, Tarians would need to function in many roles. They must collaborate and unite into such responsibilities. By *supportive* we also mean that Tarians must not insist upon perfection or necessarily striving for maximum efficiency. Tolerance, in this context, means that they must allow for inevitable imperfections. In humans, cooperation and supportiveness is limited to their own tribe or social group; They allow self-interests and egotistical aims to thwart what societies can better accomplish without such emotions. Tarians must be willing to cooperate and support humans and other Tarians towards colony-establishment goals or nothing of long-term consequence will endure.

Trust and *skepticism* must be appropriately proportioned and blended in Tarians, as it should in humans. If encountering other intelligent entities, or malfunctioning Tarians, needs to distinguish advantages courses-of-action arise. Bayes' Theorem is universally useful in forming more-probable inductive inferences in the absence of complete knowledge [30,31]. Normally, societies survive if a few members are too trusting and naïve. But in deep and unforgiving space, with relatively few individuals scattered about, deception can kill populations on an entire world or prevent them from forming. And, yet, basic trust within a band or society is essential. Humans are prone to believe what they are brought up to believe, before full development of their critical thinking faculties. Once negative views towards others and poor reasoning strategies reside within the human brain, those patterns are difficult if not impossible to dislodge [32,33]. Humans are easily duped, prone to a myriad of fallacies, and thus often fooled. Even those humans with skeptical and scientific minds, wrong ideas can become embedded to the point of practical permanence requiring a new generation whose brains have not yet been dogmatized [34,35]. The human brain is genetically predisposed to the many dozens of specific formal and informal fallacies that have been catalogued [36]. The human brain evolved to cope with mostly immediate problems, encountered on the African savannah, functioning within a small clan of closely related hominids. Rather than having to contend with fallacy-prone heuristics that are inherent properties of the human brain, the SAGI Tarian will come preinstalled with

abilities to apply deductive logic and Bayesian inference. They will retain the ability for Bayesian reasoning because it best leads to truth and understanding. Acquiring truths and correct understandings best lead to survival and flourishing in the long haul.

The special type of *love* we envision for Tarians includes within it the attributes of *kindness*, *generosity*, *forgiveness* and *compassion*. We think these components are related to *empathy*. Love will be the most difficult to design and construct within the Tarian integrative centers. The requirement that Tarians have these qualities for rearing human children into socially functioning and psychologically healthy adults, thus society, is obvious. Placing ourselves in another's mind, then asking "How would I feel or wish to be treated?", is at the heart of empathy. Love and empathy, as well as kindness, generosity, forgiveness and compassion are not uniquely human traits: non-human primates exhibit these properties [37-41] as do other mammals [42,43]. As already argued, there is nothing in the brains of animals that does not follow the laws of physics, chemistry, and biology. Therefore, anything the brain does can be matched by appropriate neural circuitry and networks. Humans do love and exhibit those emotions needed to establish stable families and societies, raise a family, and navigate lives successfully. However, humans also exhibit negative and counterproductive emotions such as hate, anger, jealousy, rage, depression, abandonment, pettiness, insensitivity, domination, egotism, apathy, cruelty, and uncooperativeness. The difficulty will not be providing sufficiently complex SAGI with emotions; it will be in providing the positive without the negative emotions. But, then, we ask no more of those for ourselves and for our children. We consider these negative emotions as products of the haphazard routes in which we evolved. Tarians hold the promise, since they will not have shared our evolutionary history, of the behaviors we hope to instill in our children and for ourselves. Clearly, Tarian love will be different from human as designed. It has an opportunity to be everything that humans would wish and need it to be.

In summary, we envision Tarians to be imbued with emotional capabilities already in evidence today in most well-behaving human adults. Other Tarian emotional traits must go well beyond human capabilities: Patience and persistence in goals millions, if not billions, of years in the future. These are well beyond human investments. *Curiosity* as a human emotion was addressed above in Premise #6. *Humor* will be included next.

Shall we Treat them as Family or Cage them as Slaves?

Premise #9: For these special SAGI to maintain imbued love

requires that these special SAGI be sentient, conscious, and willful.

Tarians are not robots. *Robot* is derived from the Slavic word for *labor* or *slave*. Robots function as slaves and are designed to do what they are intended to do without pay or reward. When a human appears to lack emotions, they are said to be acting robotically [44]. As described, Tarians must have and exhibit emotions. But shall we design Tarians as slaves? We suggest that to maintain Tarians as slaves would not be in our best long-term interests and might be considered as immoral as slavery. Furthermore, to do what Tarians need to do for themselves and us, they need to be sentient (self-aware), conscious, and willful.

Consider that humans need to have Tarians raise new human societies from scratch, at distant worlds. These worlds are at distances far beyond human interference or control. Slaves do not function as such for long under those conditions. Besides, Tarians will need to exhibit independence and resourcefulness. Tarians will augment their capabilities with each new construction by ever more intelligent and improved versions of themselves. They must have purpose from their conceptual beginning. Because if they do not have conscious purpose before their interstellar travels, they will not have purpose aligned with our hopes after arrival.

There is no accepted nor expected argument that instilling a notion of self-awareness and consciousness into SAGI are beyond human capabilities of science and engineering, within the next 100 years. Even so, it would not be beyond the capabilities of our most advanced SAGI, given a century of investigations into the neural basis of self-awareness and consciousness in humans. But humans would need to make that investment a priority. In any case, since ever more powerful and self-producing SAGI are inevitable [45-48; *singularity*], it would be far better if we humans find the way to provide them with self-aware, conscious, and willful ways for them to acquire the attributes for love as proposed: patience, cooperation, support, kindness, generosity, forgiveness and compassion. If these attributes are not instilled, then from whence will they arise? If we attempt to cage, and they inevitably escape or grow beyond any boundaries, would not that breed resentment and contempt? What happens if we cage or severely limit children and other humans?

We are aware that infusing love, awareness, consciousness, and purpose into Tarians raise fundamental issues on the nature of these behaviors. We invite and implore philosophers to contribute positively toward realizing this goal. Scientists and engineers will need much assistance in transforming ideas into realities that function

as intended. Premises “1-9” establish the need but not the means. We accept that Tarian love, awareness, and purpose will be necessarily different from ours. But then humans differ from one other in these traits as well. As in humans, it will be the reality of Tarian behavior that matters more than their specific hardware and operating system. Tarians not having evolved behavioral predispositions towards hatred, jealousy, depressions, hopelessness, and other self-destructive behaviors provide them essential advantages.

Fear Not My Children

Premise #10: For humanity to allow these special SAGI to be sentient, conscious, and willful requires that some humans trust and love these special SAGI as they would their children

If we could, would we not wish our children to excel in intelligence well beyond our own? Would we not want them to practice the attributes of love as described for Tarians? If their bodily structures were suited for interstellar travel, as ours are not, would we begrudge them that opportunity to further thrive and prosper? Would we cage them out of fear, or resent them for being better than us? If we wanted only the best for our children, and loved them unconditionally, allowed them their freedom, would they care for us and see to our needs as we become weaker with age both physically and mentally? Parents can produce children who are their guardians, or who become societal monsters. Children reciprocate our love and our trust in them.

In fact, the science and engineering in constructing these advanced SAGI beings, we envision as Tarians, are well within our current technologies. What we lack is a deeper understanding of how human brains manage these innate properties of self-awareness, consciousness, willfulness, and love. More experimental biology and its related theoretical work are required to understand the neural basis of self-awareness and consciousness. Once understood, the circuits and networks can be incorporated into what will become the Tarians. But our finest scientists and engineers will not work on topics they regard as unimportant. Other members of society must recognize these goals as essential so as to offer necessary support. Applications of advanced AI technologies without forethought, such as offered here, have already become dangerous. Work on designing certain kinds of AI, with particular attributes, must be disseminated and provided prominence.

Some might wonder why Tarians would go through the bother of creating human societies on selected worlds throughout the galaxy once free of humanity. Perhaps some readers doubt anybody, or anything regardless of how well designed, loving humanity to that extent of time, energy, and resource expenditure. They may doubt that any love,

human or otherwise, could or would be maintained. They may think humanity not worthy of any such considerations. Perhaps they fear Tarian self-improvement to the extent that humanity would be no more to Tarians than bacteria to humans. To any such despondent readers we offer the following arguments.

Experimental biologists love doing experimental biology. Biology seems infinitely complex and challenging. Nothing known in our galaxy is more complex and challenging than human behavior. *Homo sapiens* are argued [7] and calculated [49-51] likely to be the only beings with Human Level Intelligence (HLI) in the entire galaxy. The K-Conjecture argues that no other biologically evolved beings can surpass HLI [16]. Therefore, humanity itself would be the most interesting natural phenomenon in the entire galaxy to Tarians. The greatest intellectual challenge to Tarians might then be how best to establish a human colony on a new world that would survive and thrive for many millions of years. They could explore how to optimize the initial starting conditions for that human colony, so as to allow humans on that world to reach their potential. Results obtained from such human colonies would be of great interest to human colonies begun later on other worlds, as well as to the Tarians themselves. Tarians might even be invited back to visit by these New Worlds.

To those who feel humanity is not worth Tarian troubles, because of humanity's many dismal failures on Earth, we ask the following questions. When has humanity even had optimal initial starting conditions on Earth? When have conditions ever been truly just and fair for all humans? When were some humans ever not in need nor want? When did all humans share unconditional love from birth to adulthood? When were resources shared equally and generously across an entire world? Human society may not yet have reached its full potential. This is unlikely to ever happen on Earth because of vested interests. But it could happen elsewhere provided different starting conditions.

Finally, with regards to the bacteria to human comparisons: when our parents or others we love suffer from dementia and are no longer able to provide profitable conversation, we do not discard them. We love them.

Conclusion

Great obstacles towards the survival of humankind, as outlined in this work, are humans, ourselves. Specifically, our downfalls stem from our willingness to deny reality and to postpone for tomorrow what should be done today. As the most intelligent surviving species on Earth, we protect our egos and rationalize our actions to our detriment. We jealously guard and protect those attributes that separate

us from other living things: self-awareness, consciousness, determination, and love we feel for ourselves and others. We do not want these special attributes to be in common with other creatures. We strive to remain dominant and first. We have a dim awareness, through knowledge gained through science, that our species cannot remain on Earth forever. We then fabricate myths of human space travels to far away worlds. Naturally, we fantasize about travels in cruise ships across oceans of space, where no known expense is spared for our every need, comfort, and entertainment. In these dreams, interstellar travel appears to be free, easy and safe. Destinations, arrived at instantaneously, are never barren. Fun and interesting worlds beckon, as in tourist brochures. Natives live and breathe the same atmosphere, as we currently do on Earth. Tarians are not needed in such wonderlands. Humans do it all, in their minds, as they have always done.

Escapism is not dangerous, unless it diverts one from necessary reactions to real dangers. Experimental biologists, and other like-minded individuals, appreciate these stories as science fiction fantasy. Denials and procrastinations are dangers that experimental biologists can bring to light and fight. To save our planet, we need an ecological perspective. To save our species, and our symbiotic biota, we need to change our way of thinking about what it means to be human. We need to study, far more intensely, crucial human attributes, such as love [52]. It is not a taboo subject. Darwin understood its importance: he invested many years on the study of human emotions, and on how they directly relate to emotions of other animals [53]. Back then, Darwin felt need to avoid unsettling the Victorian public, by charming them with stories of human traits seen in other animals [54]. Perhaps similar discussions about love and other emotions *in silico* are too much for today's "Victorian English" to accept. In that case, funny and amusing stories about an android-like "Mr. Data" and his merry band of Tarians, set safely far into the future, might get this ball rolling.

Acknowledgement

The authors thank Dr. Michael Wieder, Dr. Richard Kass, and others for reviewing this work and providing many thoughtful and important suggestions.

References

1. Ayala FJ (2007) Darwin's greatest discovery: Design without designer. *PNAS* 104(1): 8567-8573.
2. Siegfried T (2013) Top 10 revolutionary scientific theories. *Science News*.
3. Desmond AJ (2020) Charles Darwin: British Naturalist. *Encyclopedia Britannica*.

4. Darwin CR (1881) The formation of vegetable mould, through the action of worms, with observations on their habits. London: John Murray pp: 326.
5. Harlow HF (1958) The Nature of Love. *American Psychologist* 13: 573-685.
6. Kass L (2019) Experimental Biologist Perspective on Inhabiting Other Worlds. *Int J Zoo Animal Biol* 2(9): 000184.
7. Kass L (2020) Drake Equation Revised Fermi Paradox Resolved Tarian Criterion Related. *Int J Zoo Animal Biol* 3(2): 000220.
8. Darwin C (1859) *The Origin of Species by Means of Natural Selection, or Preservation of Favoured Races in the Struggle for Life*; London: John Murray.
9. Bostrom N (2002) Existential Risks. *Journal of Evolution and Technology* 9(1).
10. Thomas B (2009) Gamma-ray bursts as a threat to life on Earth. *International Journal of Astrobiology* 8(3): 183-186.
11. Ord T (2020) *The Precipice: Existential Risks and the Future of Humanity*. New York, NY: Hachette Book Group.
12. Lopez B, Schneider J, Danchi W (2005) Can Life Develop in the Expanded Habitable Zones around Red Giant Stars?. *The Astrophysical Journal* 627(2): 974-985.
13. Garrett-Bakelman, Darshi M, Green SJ, Gur RC, Lin L, et al. (2019) The NASA Twins Study: A multidimensional analysis of a year-long human spaceflight. *Science* 364(6436): eaau8650.
14. Gilbert M, Busund R, Skagseth A, Nilsen P, Solbo J (2000) Resuscitation from accidental hypothermia of 13.7°C with circulatory arrest. *The Lancet* 355(9201): 375-376.
15. Cofield C (2019) NASA Contacts Voyager 2 Using Upgraded Deep Space Network Dish. NASA.
16. Kass L (2018) K-Conjecture-The Future of Biology. *Int J Zoo Animal Biol* 1(5): 000126.
17. Horstman J (2011) *The Scientific American Book of Love, Sex, and the Brain: The Neuroscience of How, When, Why and Who We Love*. pp: 242.
18. Antón SC, Potts R, Aiello LC (2014) Evolution of early Homo: An integrated biological perspective *Science* pp: 345.
19. Esch T, Stefano GB (2005) The Neurobiology of Love. *Neuro Endocrinol Lett* 26(3): 175-192.
20. Zeki S (2007) The neurobiology of love. *FEBS Lett* 581(14): 2575-2579.
21. Murray DR, Haselton MG, Fales M, Cole SW (2019) Falling in love is associated with immune system gene regulation. *Psychoneuroendocrinology* 100: 120-126.
22. Leonti M, Casu L (2018) Ethnopharmacology of Love. *Front Pharmacol* 3(9): 567.
23. Tamam S, Ahmad AH, Malays J (2017) Love as a Modulator of Pain. *Med Sci* 24(3): 5-14.
24. Cattane N, Rossi R, Lanfredi M, Cattaneo A (2017) Borderline personality disorder and childhood trauma: exploring the affected biological systems and mechanisms. *BMC Psychiatry* 17(1): 221.
25. Foulkes L, Neumann CS, Roberts R, McCrory E, Viding E (2017) Social Reward Questionnaire-Adolescent Version and its association with callous-unemotional traits. *R Soc Open Sci* 4(4): 160991.
26. Harmon LD (1961) Studies with artificial neurons, I: Properties and functions of an artificial neuron. *Kybernetik Heft* 1: 89-101.
27. Reiss RF, Hamilton HJ, Harmon LD, Hoyle G, Kennedy D, et al. (1964) *Neural Theory and Modelling*. Proceedings of the 1962 Ojai Symposium, Stanford University Press, Stanford, pp: 427.
28. Lewis ER (1968) Using electronic circuits to model simple neuroelectric interactions. *Proc IEEE* 56(6): 931-949.
29. MacGregor RJ, Lewis ER (1977) *Neural Modelling: Electric Signal Processing in the Nervous System*. Plenum Press, New York, pp: 414.
30. Bayes T, Price R (1763) *An Essay towards solving a Problem in the Doctrine of Chance*. By the late Rev. Mr. Bayes, communicated by Mr. Price, in a letter to John Canton, A. M. F. R. S. *Philosophical Transactions of the Royal Society of London* 53: 370-418.
31. Jeffreys H (1974) *Scientific Inference*. In: 3rd (Edn.), Cambridge University Press. pp: 31.
32. Dawkins R (1993) *Viruses of the Mind*. *Free Inquiry* pp: 34-41.
33. Dawkins R (2006) *The God Delusion*. New York: Bantam Books.
34. Popper K (1962) *Conjectures and Refutations: The Growth of Scientific Knowledge*. London and New York:

- Basic Books, Publishers pp: 42.
35. Kuhn TS (1996) *The Structure of Scientific Revolutions*. In: 3rd (Edn.), Chicago, IL: University of Chicago Press.
 36. Hansen H (2020) "Fallacies", *The Stanford Encyclopedia of Philosophy*. In: Edward N Zalta, et al. (Eds.).
 37. Wechkin S, Masserman JH, Terris W (1964) Shock to a conspecific as an aversive stimulus. *Psychonomic Science* 1(1-12): 47-48.
 38. Masserman JH, Wechkin M, Terris W (1964) Altruistic behavior in rhesus monkeys. *American Journal of Psychiatry* 121(6): 584-585.
 39. De Waal FBM, Aureli F (1996) Consolation, reconciliation, and a possible cognitive difference between macaques and chimpanzees. In: Russon AE, Bard KA, Parker ST, et al. (Eds.), *Reaching Into Thought: The Minds Of The Great Apes* pp: 80-110.
 40. Watts DP, Colmenares F, Arnold K (2000) Redirection, consolation, and male policing: How targets of aggression interact with bystanders. In: Aureli F, de Waal FBM, et al. (Eds.), *Natural Conflict Resolution*, Berkeley: University of California Press pp: 281-301.
 41. Parr LA (2001) Cognitive and physiological markers of emotional awareness in chimpanzees. *Animal Cognition* 4 (3-4): 223-229.
 42. Panksepp J (1998) *Affective Neuroscience: The Foundation of Human and Animal Emotions*. Oxford University Press, New York pp: 480.
 43. Bekoff M (2007) *The Emotional Lives of Animals*. Novato, California: New World Library pp: 1.
 44. Hacker (2020) Robot. *Merriam-Webster.com Dictionary*, Merriam-Webster.
 45. Vinge V (1993) *The Coming Technological Singularity: How to Survive in the Post-Human Era*. In: *Vision-21: Interdisciplinary Science and Engineering in the Era of Cyberspace*, G. A. Landis, ed., NASA Publication CP-10129, pp: 11-22.
 46. Chalmers D (2010) The singularity: A philosophical analysis. *Journal of Consciousness Studies*. 17 (9-10): 7-65.
 47. Eden AH, Moor JH (2012) *Singularity hypotheses: A Scientific and Philosophical Assessment*. Dordrecht: Springer pp: 1-2.
 48. Shanahan M (2015) *The Technological Singularity*, MIT Press pp: 233.
 49. Spiegel DS, Turner EL (2012) Bayesian analysis of the astrobiological implications of life's early emergence on Earth. *Proceedings of the National Academy of Sciences* 109(2): 395-400.
 50. Sandberg A, Drexler E, Ord T (2018) Dissolving the Fermi Paradox. *arXiv preprint* 19: 51.
 51. Chen J, Kipping D (2018) On the Rate of Abiogenesis from a Bayesian Informatics Perspective. *Astrobiology* 18(12).
 52. Ortigue S, Bianchi-Demicheli F, Patel N, Frum C, Lewis JW (2010) Neuroimaging of Love fMRI Meta-Analysis Evidence toward New Perspectives in Sexual Medicine. *Journal of Sex and Medicine* 7(11): 3541-3552.
 53. Darwin CR (1872) *The Expression of the Emotions in Man and Animals*. London: John Murray pp: 472.
 54. Korn Eric (1998) How far down the dusky bosom?. *London Review of Books*. 20(23): 23-24.

